Application of Variational Autoencoders in Aircraft Turbomachinery Design

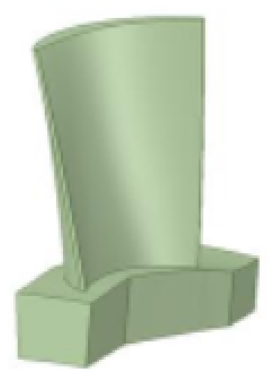
Jonathan Zalger
jzalger@stanford.edu

Motivation

Machine learning is often used in engineering for design optimization or to build surrogate models for computationally expensive engineering simulations. These applications typically involve small numbers of prespecified parameters and fail to capture higher level intuition of experienced engineers.

By including plots and images into the machine learning model, optionally ranked or classified by an engineer, we can attempt to capture more subtle engineering intution in a more intuitive manor than traditional methods.

In this project we build a variational autoencoder[1] to model transonic airflow characteristics on a NASA Rotor 37 compressor blade[2] in response to changing inlet massflow conditions.



Once the trained model is built, the flow field at new conditions can then be sampled from the posterior distribution of the latent variables given constraints on the latent manifold.

Dataset

The dataset for this project is generated via a numerical simulation technique known as computational fluid dynamics (CFD) which simulates the flow field around the rotating blade. The set of data points is generated using Latin Hypercube Sampling [3] across a range of inlet mass flow boundary conditions.

From each simulation we extract an 84 x 25 x 3 (pixels) colour image showing relative mach number contours at mid span. To form the input vector for the VAE model, we flatten the image pixels and add the corresponding mass flow condition for that point. For this project, the samples are not classified and thus the learning problem is unsupervised.

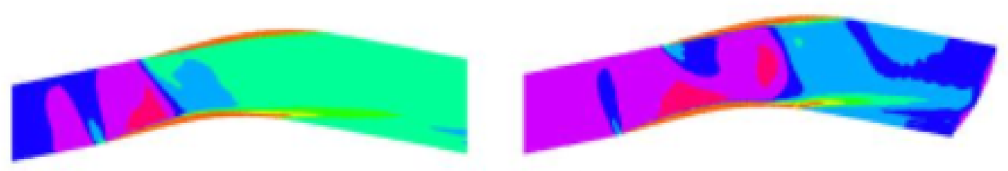


Figure 1. Example dataset images showing relative mach at mid span

Model

Variational autoencoders consist of an encoding-decoding pair of neural networks. The encoder takes the input and computes the mean and variance of a latent gaussian distribution. The decoder receives a sampled value from the latent space and generates the mean and variance of the output[1].

$$L(\theta, \phi, x^{(i)}) = \frac{1}{L}\sum_{l=1}^{L}(\log p_\theta(x^{(i)}|z^{(i,l)})) - D_{KL}(q_\phi(z|x^{(i)})||p_\theta(z))$$

The objective function of the VAE maximizes the reconstructive likelihood subject to regularization by the KL divergence of the latent space with respect to a unit gaussian distribution [1].

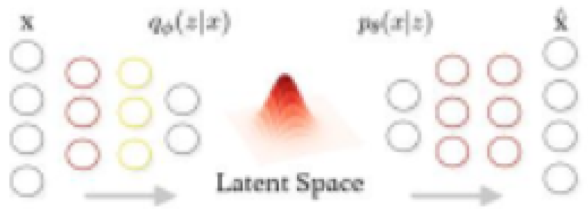


Figure 6. VAE network architecture

- 200 units/hidden layer
- 2 encoder layers with dropout
- 2 decoder layers
- Elu hidden activations [4]
- Sigmoid decoder output

Discussion

In general, the model was fairly successful in encoding a latent distribution to capture the two primary flow states modeled. A well performing model should output similar SSIM values across datasets. Note that values near 100% would indicate poor generalization and are not necessarily desired. In early error analysis, the largest issue identified was overfitting. This is evident in early latent space models (figure 4) and can be seen plotting SSIM against dataset size (figure 5). To address this, a dropout layer was added to the encoder, the image size reduced, and the datasize increased by 3600 images.

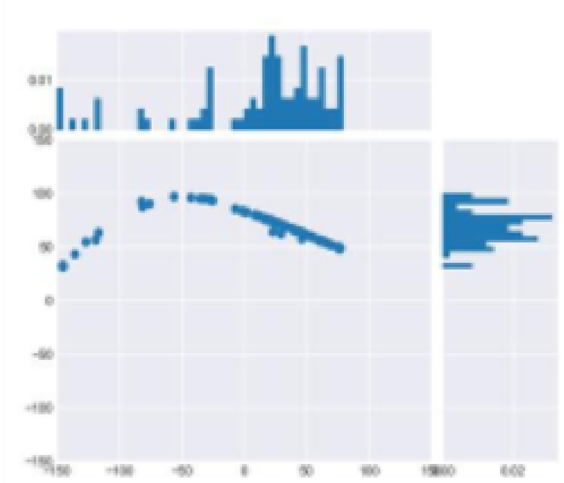


Figure 4. Latent space without dropout

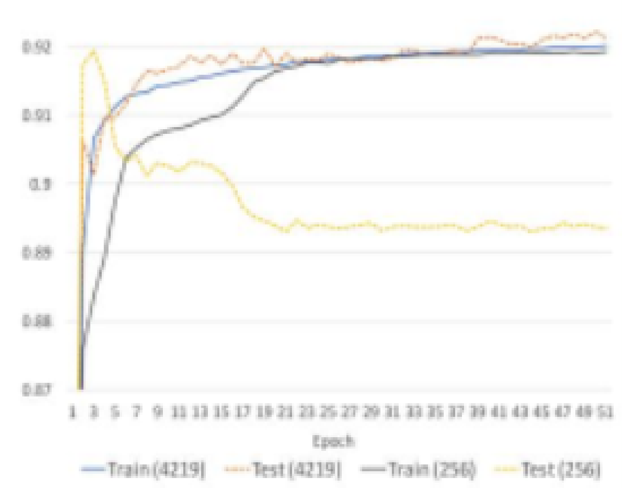


Figure 5. Test/Train SSIM by dataset size

Results

To estimate the reconstructive performance of the model we use the image structural similarity (SSIM) [6]. Table 1 illustrates the SSIM scores for the training, development, and test sets. Figure 3 shows the test samples encoded into the 2D latent space and the distribution of the latent variables. In this figure we can see two distinct states emerging corresponding to the dominant shock wave positions in the dataset. In figure 2, we show resulting images if sampled from the corresponding region of the latent space.

Table 1. Reconstructive Similarity (SSIM)

Training Set (4129 samples)	90.5 %
Development Set (498 samples)	92.2 %
Test/Validation Set (245 samples)	90.3 %

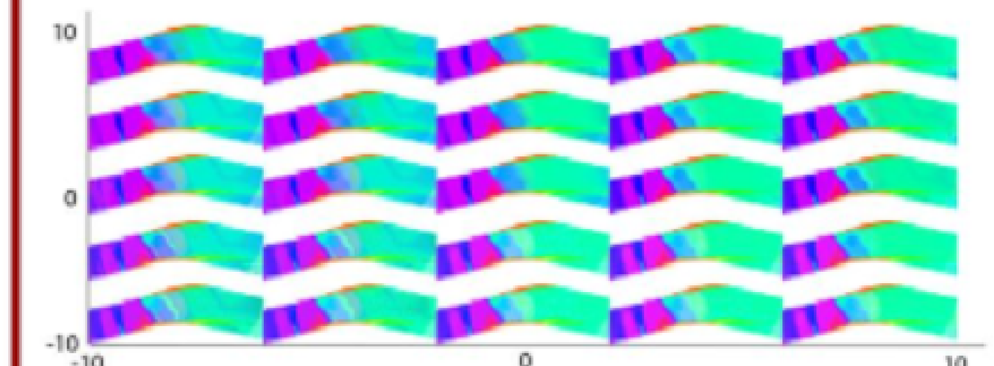


Figure 2. Latent space visualization

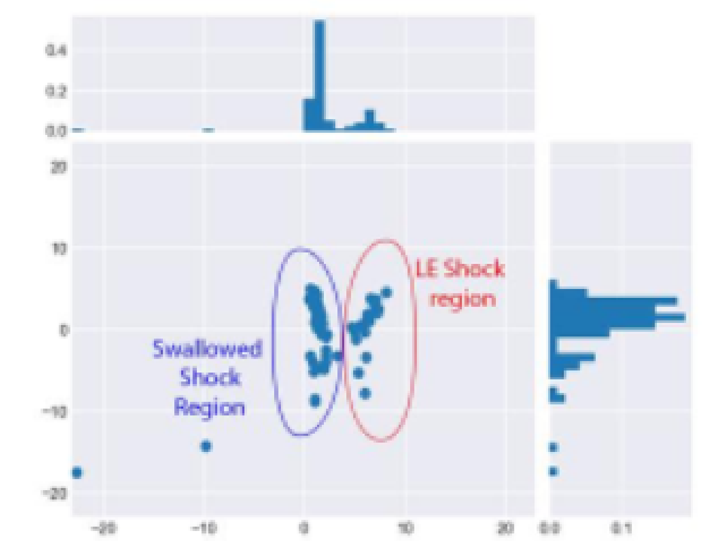


Figure 3. Encoded test samples in latent space

Future Work

The most advantageous future work involves addressing the overfitting and dataset bias issues. These can be improved immediately through expanding the dataset size, particularly in the swallowed shock state where CFD convergence issues reduce the overall sample yield. The use of convolutional layers may also improve the generalization capability [5].

There are also many other exciting applications within the aerospace field including the variation of geometric blade parameters or other boundary conditions. Exploring the capability of latent space arithmetic may also open new ways to use such models in design optimization tasks.

References

1. Diederik P Kingma and Max Welling. Auto-Encoding Variational Bayes. arXiv.org, December 2013.

2. J. D. Denton. Lessons from rotor 37. Journal of Thermal Science, 6(1):1–13, Mar 1997.

3. Jaroslav Men ˇc ık. Concise Reliability for Engineers. Concise Reliability for Engineers, (Chapter 16), April 2016.

4. ELU-Networks: Fast and Accurate CNN Learning on ImageNet. pages 1–1, December 2015.

5. Ladislav Rampasek, Daniel Hidru, Petr Smirnov, Benjamin Haibe-Kains, and Anna Goldenberg. Dr.VAE: Drug Response Variational Autoencoder. arXiv.org, June 2017.

6. D. M. Rouse, S. S. Hemami. Understanding and simplifying the structural similarity metric. IEEE. ICIP 2008.